2003

Korsgaard on the Foundations of Moral Obligation

Jason Baehr
jbaehr@lmu.edu

Recommended Citation

Baehr, Jason S. "Korsgaard on the Foundations of Moral Obligation." Journal of Value Inquiry, vol. 37, no. 4, Jan. 2003, pp. 481–491.

The Journal of Value Inquiry **37**: 481–491, 2003.

Korsgaard on the Foundations of Moral Obligation

JASON S. BAEHR
Department of Philosophy, Baylor University, P.O. Box 97273, Waco, TX 76798-7273, USA; e-mail: Jason_Baehr@baylor.edu

One of the most influential works to emerge from the recent renewal of interest in Kantian moral philosophy is Christine Korsgaard's *The Sources of Normativity*.[1] As the title suggests, Korsgaard's aim in this book is to develop an account of the foundations of moral obligation. Her basic argument is straightforward and noticeably Kantian: we must value our own humanity; thus, except on the pain of inconsistency, we must also value the humanity of others; and valuing the humanity of others amounts to being morally obligated to them in certain substantive ways.[2] Korsgaard gives considerably more attention to the second and third steps of her argument than she does to the first. However, the first step, which states that we must value our own humanity, is an equally critical part of Korsgaard's general account of moral obligation. For Korsgaard, we are obligated to value the humanity of others ultimately because we must value our own humanity. Therefore if it turns out that we need not value our humanity, Korsgaard's entire account of moral obligation is left ungrounded. What does Korsgaard mean when she claims that we must value our own humanity? How does she go about defending her claim? These questions are critical to an assessment of Korsgaard's account of moral obligation and yet very difficult to answer, mainly because Korsgaard's treatment of her value claim is brief and not very perspicuous.

1. Reasons and Practical Identities

At the heart of Korsgaard's account of moral obligation are two general claims. The first is that as human beings we need reasons in order to act. This is our lot as rational beings, or as beings that stand at a certain reflective distance from their desires and impulses:

> I desire and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act? Is

> this desire really a *reason* to act? The reflective mind cannot settle for perception or desire, not as such. It needs a *reason*. Otherwise, at least as long as it reflects, it cannot commit itself or go forward.[3]

Unlike other animals, our actions are not determined by our desires and impulses. We need reasons in order to act. The second claim concerns the source of our reasons for action. Korsgaard argues that these reasons arise from our "practical identities," which she defines as "a description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking."[4] She explains:

> You are a human being, a woman or a man, an adherent of a certain religion, a member of an ethnic group, a member of a certain profession, someone's lover or friend, and so on. And all of these identities give rise to reasons and obligations. Your reasons express your identity, your nature; your obligations spring from what that identity forbids.[5]

As human beings, we are forced to act for reasons that arise from the ways in which we value ourselves. Korsgaard clarifies the relation between these two claims as follows:

> The reflective structure of human consciousness sets us a problem. Reflective distance from our impulses makes it both possible and necessary to decide which ones we will act on: it forces us to act for reasons. At the same time, and relatedly, it forces us to have a conception of our own identity, a conception which identifies us with the source of those reasons. In this way, it makes us laws to ourselves. When an impulse – say a desire – presents itself to us, we ask whether it could be a reason. We answer that question by seeing whether the maxim of acting on it can be willed as a law by a being with the identity in question. If it can be willed as a law it is a reason, for it has an intrinsically normative structure. If it cannot be willed as a law, we must reject it, and in that case we get an obligation.[6]

Thus formulated, Korsgaard's view leaves open the possibility that we might value ourselves in ways that lead to our owing little or nothing to others. In response to this problem, Korsgaard introduces her value claim that as a matter of psychological fact, we must value ourselves as human beings. She then goes on to claim that out of consistency we must also value others as human beings, which in turn means that we have moral obligations. If she fails to establish her value claim, her account of moral obligation lacks adequate support.

2. Korsgaard's First Argument for Her Value Claim

Korsgaard defends her value claim in two main passages. In the first, she begins by observing that many of our practical identities are contingent and thus may be shed. She continues:

> What is not contingent is that you must be governed by *some* conception of your practical identity. For unless you are committed to some conception of your practical identity, you will lose your grip on yourself as having any reason to do one thing rather than another – and with it, your grip on yourself as having any reason to live and act at all. But *this* reason for conforming to your particular practical identities is not a reason that *springs from* one of those particular practical identities. It is a reason that springs from your humanity itself, from your identity simply as *a human being*, a reflective animal who needs reasons to act and to live. And so it is a reason you have only if you treat your humanity as a practical, normative, form of identity, that is, if you value yourself as a human being.[7]

Korsgaard's claim is apparently that to have a reason to take up some practical identity or other, it is necessary that we first value our own humanity.

Her reasoning in support of this claim is as follows. The reflective structure of human nature requires that we have reasons in order to act, and reasons always arise from our practical identities; therefore human nature requires that we adopt some practical identity or other in order to act. If we accept Korsgaard's view of our practical identities as the source of our reasons for action, this much of her argument is uncontroversial. She goes on to say that the fact that our own humanity requires that we adopt some practical identity or other is a reason to actually take up some practical identity or other only if we value our own humanity. Her point is that unless we value our own humanity, we will not have any reason to, as it were, satisfy the demands of our humanity: we will not have a reason to take up some practical identity or other. This claim is also uncontroversial and we may, with Korsgaard, conclude that in order to have a reason to take up some practical identity or other, we must value our own humanity.

Before considering the role this argument might play in a theory of moral obligation, it is worth pointing out a certain misleading feature of it. By "valuing our own humanity," Korsgaard means valuing ourselves as creatures who need reasons in order to act.[8] This, however, is a slightly mistaken way of putting the proper conclusion of her argument. What the argument shows is not that we must value ourselves as individuals who need reasons in order to act, but as individuals who in fact act. We must value ourselves as agents. The fact that if we fail to adopt some practical identity or other we will fail to have

any reasons for action is a reason to actually take up some practical identity or other only if we value ourselves as actors, or agents, since what is lost by not adopting a practical identity is precisely the ability to act. This means, contrary to what Korsgaard suggests, that valuing our own humanity amounts, not to valuing ourselves as creatures that need reasons in order to act, but as creatures who do act, or as agents.

Korsgaard defends her value claim mainly as a way of showing that we have certain moral obligations. Thus it is important to consider the connection between the foregoing defense of her value claim and this more general issue. Specifically, we must consider whether, in showing that to have a reason to take up some practical identity or other we must value our own humanity, Korsgaard has made any progress toward showing that we must value the humanity of others. This will apparently depend upon the extent to which it is necessary that we possess a reason to take up some practical identity or other. Korsgaard thinks that we must value the humanity of others ultimately because we must value our own humanity. But her argument shows that we must value our own humanity only in the qualified sense that this is necessary in order for us to have a reason to take up some practical identity or other. Thus, if Korsgaard hopes to provide an adequate basis for moral obligation, she must show that in some unqualified sense we need a reason to take up some practical identity or other. Otherwise, she will have failed to show that we must, in the sense intuitively required for moral obligation, value the humanity of others.

But in what sense, or under what circumstances, might a given person need a reason to adopt some practical identity or other? Perhaps Korsgaard would respond by claiming that since to act at all, the person must adopt some practical identity or other, and since he always needs reasons for adopting practical identities, he must have a reason to adopt some practical identity or other in order to act at all. This argument fails on at least two counts. It is mistaken to think that a reason is always required for the adoption of a practical identity. We commonly just find ourselves with certain practical identities. For example, through the ways in which we happen to have been raised, or the culture in which we happen to live, we often come more or less unconsciously to value ourselves in certain ways, which provide us with reasons for action. This shows that we can come to have reasons for action without having a reason of the sort in question. As well, we take up many of our practical identities on the basis of other ways in which we value ourselves. For instance, if a person values herself as a successful scholar and teacher, she will also have reason to value herself as a person who is thorough and careful in doing research, fair-minded in the evaluation of the views of others, an effective communicator of ideas, and concerned about the welfare of her students. This shows that a person can have reasons for taking up practical identities without having a reason to take up some practical identity or other, in Korsgaard's

sense. For both reasons it is mistaken to think that in order to act at all we must value ourselves as human beings. This is true only if acting in general requires having a reason to take up some practical identity or other, which it does not.

It may be that if a person called into question all of his practical identities, doubting whether he should take up any practical identity at all, then he would need to value his own humanity, since in that case, if he were to act at all, he apparently would need a reason to take up some practical identity or other. But this is a highly unlikely scenario. Again, we value ourselves in many ways and for many reasons that do not require that we even think about our own humanity. Thus we may reject the claim that to have or adopt practical identities we must have a reason to adopt some practical identity or other. John Skorupski draws a similar conclusion with regard to Korsgaard's argument. About her claim that we must value our own humanity in order to have a reason of the sort in question, he says, "Suppose I have *this* reason for conforming to my particular personal identities only if I treat my humanity as a practical, normative identity. That wouldn't show what needs to be shown, viz. that to adopt particular practical identities I must identify with my humanity."[9] Skorupski's point is that it does not follow from the fact that we must value our own humanity in order to have a reason to take up some practical identity or other that we must value our own humanity in order to take up some particular practical identity. This again is because taking up a particular practical identity does not require the possession of any such reason.

Korsgaard's first argument for the value claim fails to provide the basis for a plausible conception of moral obligation. It fails to show in any strong sense that we must value our own humanity. Thus, her broader argument to the effect that we must value the humanity of others because we must value our own humanity does not get off the ground.

3. Korsgaard's Second Argument for Her Value Claim

Korsgaard in fact advocates a stronger version of her value claim. She is not content with the claim that to have a reason to adopt some practical identity or other, we must value our own humanity. She thinks that we must value our own humanity if we are to act at all. Korsgaard defends this stronger version of her value claim in a passage that begins with a reiteration of the point that each of us is "an animal who needs a practical conception of her own identity, a conception of who she is which is normative for her."[10] She goes on to say:

> It is because we are such animals that our practical identities are normative for us, and, once you see this, you must take this more fundamental

> identity, being such an animal, to be normative as well. You must value your humanity if you are to value anything at all.
>
> Why? Because now that you see that your need to have a normative conception of yourself comes from your human identity, you can query the importance of that identity. Your humanity requires you to conform to some of your practical identities, and you can question this requirement as you do any other. Does it really matter whether we act as our humanity requires, whether we find some ways of identifying ourselves and stand by them? But in this case you have no option but to say yes. Since you are human you *must* take something to be normative, that is, some conception of practical identity must be normative for you. If you had no normative conception of your identity, you could have no reasons for action, and because your consciousness is reflective, you could then not act at all. Since you cannot act without reasons and your humanity is the source of your reasons, you must value your humanity if you are to act at all.[11]

As the final sentence of this passage indicates, Korsgaard attempts to draw a close connection between the claim that humanity is the source of our reasons for action and the claim that we must value our humanity. However, the precise nature of this connection is unclear. From the fact that "you cannot act without reasons and your humanity is the source of your reasons" it does not follow that "you must value your humanity if you are to act at all."[12]

For Korsgaard, the reflective structure of humanity requires that we have practical identities that give rise to practical reasons. Our humanity is the source of our practical reasons in the sense that it is what requires that we have these reasons. But it remains to be determined whether we must therefore value our humanity. A person's poor eyesight might require that she wear glasses, but to wear glasses, she need not value her poor eyesight. In his review of *Sources of Normativity*, Allan Gibbard raises a similar point, using an example of an ascetic: "As an ascetic . . . I might value being someone who resists the cravings of the flesh. I can't be someone who does that without having cravings of the flesh; am I then committed to valuing having these cravings?"[13] Rachel Cohon also raises a nearly identical version of this objection in her treatment of Korsgaard's value claim.[14] Thus it appears that the inference from the claim that our humanity requires that we act for reasons to the claim that we must then value our humanity is highly questionable.

There are at least two further ways of understanding the relation between these two claims. First, Korsgaard might mean that unless we in some sense value our own humanity, we will lack any reason to satisfy its demands. We will lack any reason to take up some practical identity or other. But, Korsgaard might claim, without such a reason we will not have any practical identities at all, and since the possession of a practical identity is a necessary condition for action, we will be unable to act at all. The problem with this version of the

argument is that it rests upon the same faulty assumption involved in Korsgaard's previous argument for her value claim. It assumes that we must have a reason to take up some practical identity or other in order to have a particular practical identity. This is mistaken, however, since we might have some particular practical identity either somewhat accidentally, for example, as a result of our upbringing, or as a result of the demands of one of our other practical identities. Thus, as long as we have practical identities, we have reasons for action, and to have practical identities we need not be in possession of a reason of the sort in question.

However, there remains at least one other way of understanding Korsgaard's argument which may have more promise. Her argument might be interpreted as exhibiting the following transcendental structure, which is consistent with the way that she at times describes her argument and with the generally Kantian character of her overall moral theory.[15] It is an indisputable fact that we sometimes act. Acting requires the possession of practical reasons, which, Korsgaard holds, arise from our normative self-conceptions or practical identities. However, these self-conceptions are normative only if we value ourselves as beings who act, or as agents. If we did not already value ourselves as agents, or as beings that do certain things, then the fact that a given practical identity licenses or forbids a particular action would not give us reason to do or refrain from doing what is licensed or forbidden by that practical identity. But our practical identities do give us reasons for action; they clearly are normative for us. Therefore we must value ourselves as agents, and for Korsgaard this is tantamount to valuing our own humanity.

This does appear to reveal a sense in which to act at all we must value our own humanity. It shows that valuing our humanity in some sense is a necessary condition for the very possibility action. If we failed to value ourselves as agents, our practical identities would fail to be normative for us, in which case we would lack reasons for action, and hence be unable to act. But since we do act, we must have reasons for action, which shows that our practical identities are normative for us, and thus that we value ourselves as agents. Cohon also proposes a transcendental rendering of Korsgaard's argument. She claims that we must interpret Korsgaard as claiming that "if we value anything at all, we *already* value our reflectivity, whether or not we have performed any process of reflection or act of endorsement. Our valuing of our reflective selves must be *implied* by our valuing of anything else."[16] If this way of understanding Korsgaard's argument is correct, then at a certain level she does appear to offer a successful defense of her value claim.

The critical question, however, is whether this defense of her value claim can provide an adequate foundation for Korsgaard's account of moral obligation, since she attempts to ground our obligations to others in the way that we must value ourselves. To evaluate Korsgaard's view in this regard, we must attempt to clarify the sense in which we are required to value our humanity

according to the forgoing argument. More importantly, we must attempt to specify what we might owe to others if it were to turn out that we must value their humanity in the same sense that we must value our own humanity.

4. Korsgaard's Value Claim and Moral Obligation

Given that Korsgaard takes "humanity" to be synonymous with "rational agency" or our capacity to act for reasons, valuing the humanity of others in the same way that she has managed to show that we must value our own humanity places the following constraint upon our actions: it requires that we always act in such a way that we allow others to have reasons and to act for them. It requires that we always allow others to have practical identities, or to value themselves in certain ways, since we must value our own humanity in a manner that is consistent with our having reasons for action. If we were not to value our humanity in any sense, none of our practical identities would be normative for us, and this would leave us without any reasons for action and hence unable to act at all. But since we do act, we must value our humanity at least to the extent in question. Assuming, then, that we must value the humanity of others in the same way that we must value our own humanity, it follows that what we owe to others is to treat them in a manner consistent with their having reasons for action, or valuing themselves in certain ways. While this might initially seem like a promising and deeply Kantian conception of moral obligation, closer inspection reveals that it has very serious limitations.

The problem is that on Korsgaard's view, there appear to be only minimal restrictions on the quality of the reasons upon which we must permit others to act. As long as we treat others in a way that does not prevent them from valuing themselves in a certain way, and hence from having certain reasons for action, we apparently will be fulfilling our moral obligations to them. We will be valuing their humanity in the way that Korsgaard has shown that we must value our own humanity. Valuing the humanity of others in this sense does require that we not go about killing other persons, or so controlling or manipulating them that we prevent them from acting at all. For instance, if Smith is contemplating doing something to Jones that would, from Jones's perspective, make Jones's life unworthy of living, such an action would be ruled morally impermissible on Korsgaard's view because it is inconsistent with allowing Jones to have reasons for action. However, the important point is that Korsgaard's view might not prevent Smith from still doing considerable harm to Jones, and from treating him in putatively immoral ways. For example, Smith might demand of Jones "either your money or your life," and as long as Jones values himself as alive, Smith would still be permitting him to exercise his rational agency, and would thereby apparently not be doing anything morally wrong on Korsgaard's view.

Korsgaard might reply by saying that this objection fails to fully appreciate all that is involved with allowing others to exercise their rational agency in the relevant sense. She might say that it requires not merely that we permit others to act in accordance with some practical identity or other, but that we permit them to act in accordance with all of their practical identities. It requires that we not violate or act against any of the ways in which they value themselves. In this case, while Smith's mugging Jones would permit Jones to act in accordance with his practical identity as a living person, it would prevent him from acting in accordance with what is likely to be another of his practical identities: his identity as someone who has control over his own assets. Assuming this is correct, Korsgaard could then accommodate the intuition that Smith's actions are immoral.

But this does not leave Korsgaard's view without problems. Suppose that Jones, having less than a week to live and no one to leave his assets to, does not value himself as being in full control of these assets. From Jones's point of view, forfeiting his wallet to Smith is therefore likely to be of little concern. On Korsgaard's view, Smith's actions would then be morally justified, since Smith would not be violating any of Jones's practical identities. But this is a mistaken assessment of the situation. Even if it is not particularly important to Jones that he not be mugged by Smith, Smith still acts immorally by mugging Jones. Other similar cases can be advanced in support of this objection to Korsgaard's view. A woman whose husband routinely inflicts physical and emotional violence upon her might happen not to value herself in ways that conflict with this treatment. She might deeply value herself as her husband's property and therefore not take serious issue with her husband's actions toward her. In that case, her husband's behavior would be morally impermissible even though it does not involve acting against any of the wife's practical identities.

Cases like this reveal a serious problem for Korsgaard's view. The problem can be stated in two different ways. First, Korsgaard is apparently unable to place adequate constraints on what count as reasons for action. The abusive husband's impulse to do violence to his wife clearly should not count as a reason for his actual violence. Yet as long as the wife does not value herself in ways that forbid this kind of treatment, Korsgaard does not have a way of resisting this conclusion. Another way of putting the problem is that, since Korsgaard thinks that reasons for action always derive from practical identities, she cannot place sufficient restraints on the adoption of practical identities. If she could show that reason or the moral law forbid the husband from taking his conception of himself as abusive to be a source of reasons for action regardless of what self-conceptions his wife happens to have, then she would have a way of forbidding the husband's actions. We have seen, however, that Korsgaard has no way of doing this. She succeeds at showing that in a certain respect we must value our own humanity. But even if we grant

that we must also value the humanity of others in this same respect, this obligation does not bring with it the normative weight that Korsgaard would like it to. It demands that we permit one another to self-identify in ways that give us our own respective reasons for action, and it requires that we not act against those reasons. But these constraints by themselves are far too weak to support a plausible account of moral obligation.

The temptation at this point might be to think that valuing our own humanity must require more than has been suggested and that consequently we must also owe more to others. This is indeed Korsgaard's position. She moves swiftly from the arguments we have considered to the conclusion that "to value yourself as a human being is to have a moral identity, as the Enlightenment understood it."[17] Similarly, she comments: "It follows from this argument that human beings are valuable. Enlightenment morality is true."[18] Korsgaard clearly thinks that she has established her value claim in a considerably stronger sense than she really has. The issue is not what Korsgaard or we might typically think of as being involved with valuing our own humanity or the humanity of others, but instead with the sense of her value claim that her arguments substantiate. They substantiate only a modest version of this claim, which in turn results in an anemic conception of moral obligation.

The main culprit behind this problem with Korsgaard's view is what Skorupski calls her "radical voluntarism" about practical reasons.[19] For Korsgaard all normativity arises from our practical identities. We have reasons for action just because we value ourselves in certain ways. This, however, presents a major problem when her view is confronted with cases like that of the abusive husband. If, as examples like this suggest, further constraints need to be imposed on what count as reasons for action or on the sorts of practical identities we can adopt, such constraints must somehow arise from certain facts about how we must value ourselves. The problem is that it is unclear what such constraints might amount to on Korsgaard's view or what she could say about how we must value ourselves that might in turn lead to our having substantial moral obligations. But without a solution to this problem, Korsgaard appears unable to make the needed amendments to her view, and hence is left with what is at best an objectionably thin account of moral obligation.[20]

Notes

1. Christine Korsgaard, *The Sources of Normativity* (Cambridge, England: Cambridge University Press, 1996).
2. Ibid., p. 121.
3. Ibid., p. 93.
4. Ibid., p. 101.
5. Ibid.

6. Ibid., p. 113.
7. Ibid., p. 121.
8. Ibid.
9. John Skorupski, "Rescuing Moral Obligation," *European Journal of Philosophy* 6(3), (1998), p. 350.
10. Korsgaard, op. cit., p. 123.
11. Ibid.
12. Ibid.
13. Allan Gibbard, "Morality as Consistency in Living: Korsgaard's Kantian Lectures," *Ethics* 110, (Oct. 1999), p. 154.
14. Rachel Cohon, "The Roots of Reasons," *The Philosophical Review* 109, (Jan. 2000), pp. 75–76.
15. See Korsgaard, op. cit., pp. 123–124.
16. Cohon, op. cit., p. 83.
17. Korsgaard, op. cit., p. 121.
18. Ibid., p. 123.
19. Skorupski, op. cit., p. 346.
20. I am grateful to Angela Smith and to an anonymous referee for *The Journal of Value Inquiry* for helpful comments on an earlier draft this paper.